УДК 519.25:681.5

А.П. Сарычев
Институт технической механики НАН Украины и ГКА Украины
Украина, 49005, г. Днепропетровск, ул. Лешко-Попеля, 15

ЛИНЕЙНАЯ АВТОРЕГРЕССИЯ НА ОСНОВЕ МЕТОДА ГРУППОВОГО УЧЕТА АРГУМЕНТОВ В УСЛОВИЯХ КВАЗИПОВТОРНЫХ НАБЛЮДЕНИЙ

A.P. Sarychev
The Institute of Technical Mechanics of NAS of Ukraine and SSA Ukraine
Ukraine, 49005, Dnipropetrovs'k, Leshko-Popel av., 15

LINEAR AUTOREGRESSION BASED ON THE GROUP METHOD OF DATA HANDLING IN CONDITIONS OF QUASIREPEATED OBSERVATIONS

О.П. Саричев
Інститут технічної механіки НАН України і ДКА України
Україна, 49005, Україна, м. Дніпропетровськ, вул. Лешко-Попеля, 15

ЛІНІЙНА АВТОРЕГРЕСІЯ НА ОСНОВІ МЕТОДУ ГРУПОВОГО УРАХУВАННЯ АРГУМЕНТІВ В УМОВАХ КВАЗІПОВТОРНИХ СПОСТЕРЕЖЕНЬ

Для моделирования в классе авторегрессионных уравнений разработан критерий регулярности МГУА с разбиением наблюдений на обучающие и проверочные подвыборки в условиях квазиповторных наблюдений. Доказано существование оптимального множества регрессоров. Получено условие редукции оптимальной авторегрессионной модели, которое зависит от параметров авторегрессионной модели и объёмов выборок.
Ключевые слова: неопределенность по составу регрессоров, критерий регулярности.

For modelling in a class of autoregression equations, the criterion of regularity of the GMDH with dividing of observations on training and testing subsamples in conditions of quasirepeated observations is offered. It is proved, that the optimum set of regressors exists. The condition of a reduction of the optimum autoregression model is obtained. This condition depends on parameters of autoregression model and volumes of samples.
Key words: uncertainty on structure of regressors, criterion of regulatory.

Для моделювання в класі авторегресійних рівнянь розроблено критерій регулярності МГУА з розбиттям спостережень на навчальні й перевірні підвибірки в умовах квазіповторних спостережень. Доведено існування оптимальної множини регресорів. Встановлено умову редукції оптимальної авторегресійної моделі, що залежить від параметрів авторегресійної моделі та обсягів вибірок.
Ключові слова: невизначеність за складом регресорів, критерій регулярності.

Определение порядка авторегрессионной модели в условиях структурной неопределённости по количеству и составу регрессоров является важной задачей в теории идентификации, и для её решения существуют различные подходы [1–7]. Эта задача является одним из объектов исследования в методе группового учета аргументов (МГУА) [8–12], который разработал академик НАН Украины А.Г. Ивахненко. Подход основан на разбиении выборки наблюдений на обучающую и проверочную части: на обучающей выборке оцениваются коэффициенты модели, а на проверочной оценивается качество модели. В

соответствии с принципами моделирования в МГУА для того, чтобы построить авторегрессионное уравнение оптимальной сложности, необходимо: а) указать метод оценивания коэффициентов в авторегрессионном уравнении; б) задать алгоритм генерирования авторегрессионных уравнений (структур моделей); в) разработать внешний критерий для оценки качества перебираемых структур; г) исследовать поведение математического ожидания внешнего критерия в зависимости от состава регрессоров; д) доказать существование авторегрессионной модели оптимальной сложности. При моделировании в классе авторегрессионных уравнений в МГУА традиционно применяется критерий регулярности: обучающая выборка формируется первыми $n(A)$ наблюдениями временного ряда, проверочная выборка последующими $n(B)$ наблюдениями, причём выполняется $n(A)+n(B)=n$, где n – объём исходной выборки. Применение критерия регулярности в таком виде требует контроля, поскольку динамические свойства объекта могут проявляться неодинаково на разных фазах переходных процессов.

В данной работе предлагается рассчитывать критерий регулярности в так называемой схеме квазиповторных наблюдений, которая возможна в условиях активного эксперимента. В этой схеме обучающая (A) и проверочная (B) выборки получены особым способом как пара реализаций функционирования объекта с близкими начальными условиями, качественно одинаковым характером переходных процессов и близкими состояниями в конечные моменты времени.

Априорные предположения об объекте

Пусть функционирование динамического объекта подчиняется закону в виде авторегрессионного уравнения

$$\begin{pmatrix} \overset{*}{x}_1 \\ \overset{*}{x}_2 \\ \vdots \\ \overset{*}{x}_i \\ \vdots \\ \overset{*}{x}_n \end{pmatrix} = \begin{bmatrix} \overset{*}{x}_0 & \overset{*}{x}_{-1} & \cdots & \overset{*}{x}_{1-p} \\ \overset{*}{x}_1 & \overset{*}{x}_0 & \cdots & \overset{*}{x}_{2-p} \\ \vdots & \vdots & \ddots & \vdots \\ \overset{*}{x}_{i-1} & \overset{*}{x}_{i-2} & \cdots & \overset{*}{x}_{i-p} \\ \vdots & \vdots & \ddots & \vdots \\ \overset{*}{x}_{n-1} & \overset{*}{x}_{n-2} & \cdots & \overset{*}{x}_{n-p} \end{bmatrix} \begin{pmatrix} \overset{o}{\theta}_1 \\ \overset{o}{\theta}_2 \\ \vdots \\ \overset{o}{\theta}_p \end{pmatrix} + \begin{pmatrix} \zeta_0 \\ \zeta_1 \\ \vdots \\ \zeta_{i-1} \\ \vdots \\ \zeta_{n-1} \end{pmatrix}, \quad (1)$$

которое в матрично-векторной форме можно записать в виде:

$$\overset{*}{\mathbf{x}} = \overset{*}{\mathbf{Z}}(p)\overset{o}{\boldsymbol{\theta}} + \boldsymbol{\zeta}(-1), \quad (2)$$

где $\overset{*}{\mathbf{x}}$ – ненаблюдаемый $(n \times 1)$-вектор значений выходной переменной объекта в дискретные моменты времени $t = t_i$, $i = 1,2,...,n$; n – общее число наблюдений; p – число предыдущих значений выходной переменной, которые влияют на её текущее значение; $\overset{*}{\mathbf{Z}}(p)$ – $(n \times p)$-матрица p предыдущих ненаблюдаемых значений переменной, в обозначении этой матрицы p означает, что в (1)–(2) при формировании величины $\overset{*}{x}_i$ участвуют величины $(x_{i-1}, x_{i-2},..., x_{i-p})$; $\overset{o}{\boldsymbol{\theta}}$ – $(p \times 1)$-вектор неизвестных детерминированных коэффициентов; $\boldsymbol{\zeta}(-1)$ – ненаблюдаемый

случайный $(n\times 1)$-вектор, в обозначении которого ”–1” означает, что в (1)–(2) при формировании величины x_i аддитивно участвует величина ζ_{i-1}.

В (1) – (2) предполагается, что в формировании текущего значения выходной переменной участвуют все p её предыдущих значений. В общем случае не все её предыдущие значения могут участвовать в этом формировании. Для записи моделей в общем случае введем структурные матрицы, смысл которых покажем на примере. Пусть на текущее значение выходной переменной влияют первое, второе и четвертое её предыдущих значения из максимально заданного возможного числа влияющих предыдущих значений $p=5$. Тогда вместо матрицы $\overset{*}{\mathbf{Z}}(p)$ в (1)–(2) следует записать произведение матриц

$$\overset{*}{\mathbf{Z}}(p)\mathbf{S}=\begin{bmatrix}\overset{*}{x}_0 & \overset{*}{x}_{-1} & \overset{*}{x}_{-2} & \overset{*}{x}_{-3} & \overset{*}{x}_{-4}\\ \overset{*}{x}_1 & \overset{*}{x}_0 & \overset{*}{x}_{-1} & \overset{*}{x}_{-2} & \overset{*}{x}_{-3}\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ \overset{*}{x}_{i-1} & \overset{*}{x}_{i-2} & \overset{*}{x}_{i-3} & \overset{*}{x}_{i-4} & \overset{*}{x}_{i-5}\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ \overset{*}{x}_{n-1} & \overset{*}{x}_{n-2} & \overset{*}{x}_{n-3} & \overset{*}{x}_{n-4} & \overset{*}{x}_{n-5}\end{bmatrix}\times\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\\0&0&1\\0&0&0\end{bmatrix}=\begin{bmatrix}\overset{*}{x}_0 & \overset{*}{x}_{-1} & \overset{*}{x}_{-3}\\ \overset{*}{x}_1 & \overset{*}{x}_0 & \overset{*}{x}_{-2}\\ \vdots & \vdots & \vdots\\ \overset{*}{x}_{i-1} & \overset{*}{x}_{i-2} & \overset{*}{x}_{i-4}\\ \vdots & \vdots & \vdots\\ \overset{*}{x}_{n-1} & \overset{*}{x}_{n-2} & \overset{*}{x}_{n-4}\end{bmatrix}, \qquad (3)$$

где (5×3)-матрица $\mathbf{S}$ представляет собой структурную матрицу, отражающую влияние первого, второго и четвертого предыдущих значений переменной на её текущее значение. Априорная информация о значении p и о том, какие именно предыдущие значения определяют текущее значение переменной в законе функционирования объекта (1)–(2), представляется структурной $(p\times m)$-матрицей $\mathbf{S}$ (m – число столбцов в матрице $\mathbf{S}$, которое равно числу неизвестных коэффициентов в модели). Будем пока предполагать, что эта структурная матрица задана.

С учетом введенной структурной матрицы закон функционирования (2) для общего случая формирования выходной переменной запишем в виде

$$\overset{*}{\mathbf{x}}=\overset{*}{\mathbf{Z}}(p)\,\mathbf{S}\,\overset{\mathrm{o}}{\boldsymbol{\theta}}+\zeta(-1)=\overset{=}{\mathbf{x}}+\zeta(-1), \qquad (4)$$

где $\overset{\mathrm{o}}{\boldsymbol{\theta}}$ – $(m\times 1)$-вектор неизвестных детерминированных коэффициентов; $\overset{=}{\mathbf{x}}$ – ненаблюдаемая составляющая $(n\times 1)$-вектора значений выходной переменной.

Пусть для наблюдений выходной переменной объекта выполняется

$$x_i=\overset{*}{x}_i+\varepsilon_i,\quad i=1,2,\ldots,n, \qquad (5)$$

где x_i – наблюдаемое значение переменной в момент времени $t=t_i$, $\overset{*}{x}_i$ – ненаблюдаемое значение; ε_i – случайная ненаблюдаемая ошибка измерения переменной.

Запишем с учетом (5) модель наблюдения объекта в векторной форме

$$\mathbf{x}=\overset{*}{\mathbf{x}}+\boldsymbol{\varepsilon}. \qquad (6)$$

Пусть относительно $\zeta(-1)$ выполнено

$$E\{\zeta(-1)\}=\mathbf{0}_n,\quad E\{\zeta(-1)\zeta^{\mathrm{T}}(-1)\}=\sigma_\zeta\cdot\mathbf{I}_n. \qquad (7)$$

где $E\{\cdot\}$ – знак математического ожидания по возможным реализациям случайного вектора $\boldsymbol{\zeta}(-1)$; $\mathbf{0}_n$ – $(n\times 1)$-вектор, состоящий из нулей; σ_ζ – дисперсия величины $\zeta_i(-1)$, $i=1,2,...,n$, ограниченная величина; $\mathbf{I}_n$ – единичная $(n\times n)$-матрица.

Пусть относительно $\boldsymbol{\varepsilon}$ выполнено

$$E\{\boldsymbol{\varepsilon}\}=\mathbf{0}_n,\quad E\{\boldsymbol{\varepsilon}\,\boldsymbol{\varepsilon}^{\mathrm{T}}\}=\sigma_\varepsilon\cdot\mathbf{I}_n, \tag{8}$$

где $E\{\cdot\}$ – знак математического ожидания по возможным реализациям случайных векторов $\boldsymbol{\varepsilon}$; σ_ε – дисперсия величины ε_i, $i=1,2,...,n$, ограниченная величина.

Пусть в результате наблюдения в моменты времени $t=t_i$, $i=1-2p,2-2p,\dots,0,1,2,\dots,n$, получен $(n+2p)$-вектор значений выходной переменной объекта

$$\left(x_{1-2p},x_{2-2p},\dots,x_0,x_1,x_2,\dots,x_n\right)^{\mathrm{T}}=\begin{pmatrix}\mathbf{x}(0)\\ \hdashline \mathbf{x}\end{pmatrix}, \tag{9}$$

где $(2p\times 1)$-вектор $\mathbf{x}(0)$ будет использован в качестве начальных условий.

Пусть заданы: 1) p – число предыдущих значений выходной переменной, которые влияют на её текущее значение; 2) структурная $(p\times m)$-матрица $\mathbf{S}$, которая определяет, какие именно предыдущие значения переменной определяют её текущее значение.

Для оценивания неизвестных коэффициентов $\overset{o}{\boldsymbol{\theta}}$ по наблюдениям объекта (9) используем результаты [13–14], где разработаны итерационные процедуры параметрической идентификации для системы авторегрессионных уравнений.

Оценивание коэффициентов в авторегрессионном уравнении

Для $(n\times 1)$-вектора $\mathbf{x}$ выполняется

$$\mathbf{x}=\overset{=}{\mathbf{Z}}(p)\,\mathbf{S}\,\overset{o}{\boldsymbol{\theta}}+\boldsymbol{\xi}(k), \tag{10}$$

где матрица $\overset{=}{\mathbf{Z}}(p)$ – $(n\times p)$-матрица ненаблюдаемых значений переменной объекта, по своей структуре она аналогична матрице $\overset{*}{\mathbf{Z}}(p)$ в (1)–(2):

$$\overset{=}{\mathbf{Z}}(p)=\begin{bmatrix}\overset{=}{x}_0 & \overset{=}{x}_{-1} & \cdots & \overset{=}{x}_{1-p}\\ \overset{=}{x}_1 & \overset{=}{x}_0 & \cdots & \overset{=}{x}_{2-p}\\ \vdots & \vdots & \ddots & \vdots\\ \overset{=}{x}_{n-1} & \overset{=}{x}_{n-2} & \cdots & \overset{=}{x}_{n-p}\end{bmatrix}, \tag{11}$$

$\boldsymbol{\xi}$ – случайный $(n\times 1)$-вектор с нулевым математическим ожиданием

$$\boldsymbol{\xi}=\boldsymbol{\varepsilon}+\boldsymbol{\Gamma}(-2;Z)\,\mathbf{S}\,\overset{o}{\boldsymbol{\theta}}+\boldsymbol{\zeta}(-1),\quad E\{\boldsymbol{\xi}\}=\mathbf{0}_n; \tag{12}$$

$\boldsymbol{\Gamma}(-2;Z)$ – $(n\times p)$-матрица ненаблюдаемых случайных величин

$$\mathbf{\Gamma}(-2;Z)=\begin{bmatrix}\zeta_{-1} & \zeta_{-2} & \cdots & \zeta_{-p}\\ \zeta_{0} & \zeta_{-1} & \cdots & \zeta_{1-p}\\ \vdots & \vdots & \ddots & \vdots\\ \zeta_{n-2} & \zeta_{n-3} & \cdots & \zeta_{n-1-p}\end{bmatrix}. \qquad (13)$$

Введем обозначения

$$\mathbf{y}=\mathbf{x}\,,\quad \mathbf{R}=\overset{=}{\mathbf{Z}}(p)\mathbf{S}\,, \qquad (14)$$

где $\mathbf{R}$ – $(n\times m)$-матрица регрессоров для выходной переменной.

Учитывая (14), регрессионную модель (10) запишем в виде

$$\mathbf{y}=\mathbf{R}\overset{o}{\boldsymbol{\theta}}+\boldsymbol{\xi}=\overset{o}{\mathbf{y}}+\boldsymbol{\xi}, \qquad (15)$$

где $\mathbf{y}$ – $(n\times 1)$-вектор наблюдаемых зашумленных значений выходной переменной; $\overset{o}{\mathbf{y}}$ – $(n\times 1)$-вектор ненаблюдаемых значений; $\overset{o}{\boldsymbol{\theta}}$ – $(m\times 1)$-вектор неизвестных коэффициентов; $\boldsymbol{\xi}$ – $(n\times 1)$-вектор ненаблюдаемых случайных аддитивных составляющих.

Согласно [13], для оценки коэффициентов $\overset{o}{\boldsymbol{\theta}}$ выполняется

$$\hat{\mathbf{d}}=\mathbf{C}\,\mathbf{y}\,, \qquad (16)$$

где для $(m\times n)$-матрицы $\mathbf{C}$ выполняется

$$\mathbf{C}=(\mathbf{R}^{\mathrm{T}}\mathbf{\Sigma}_{\xi}^{-1}\mathbf{R})^{-1}\mathbf{R}^{\mathrm{T}}\mathbf{\Sigma}_{\xi}^{-1}\,, \qquad (17)$$

а $\mathbf{\Sigma}_{\xi}$ – ковариационная $(n\times n)$-матрица введённого в (15) $(n\times 1)$-вектора ненаблюдаемых аддитивных случайных составляющих $\boldsymbol{\xi}$.

Для ковариационной матрицы $\mathbf{\Sigma}_{\xi}$ выполняется [13]:

$$\mathbf{\Sigma}_{\xi}=\sigma_{\varepsilon}\cdot\mathbf{I}_n+\mathbf{\Psi}+\sigma_{\zeta}\cdot\mathbf{I}_n\,, \qquad (18)$$

где σ_{ε} – дисперсия в модели наблюдения, введённая в (8); σ_{ζ} – дисперсия в модели функционирования, введённая в (7); $\mathbf{\Psi}$ – $(n\times n)$-матрица имеет вид

$$\mathbf{\Psi}=\begin{bmatrix}\psi(0) & \psi(+1) & \cdots & \psi(p-1) & 0 & \cdots & 0 & 0\\ \psi(-1) & \psi(0) & \cdots & \psi(p-2) & \psi(p-1) & \cdots & 0 & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots\\ \psi(1-p) & \psi(2-p) & \cdots & \psi(0) & \psi(+1) & \cdots & 0 & 0\\ 0 & \psi(1-p) & \cdots & \psi(-1) & \psi(0) & \cdots & 0 & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & \cdots & 0 & 0 & \cdots & \psi(0) & \psi(+1)\\ 0 & 0 & \cdots & 0 & 0 & \cdots & \psi(-1) & \psi(0)\end{bmatrix}. \qquad (19)$$

В (19) величины $\psi(\Delta)$, $\Delta=-p+1,-p+2,...,p-2,p-1$, определяются по формулам

$$\psi(\Delta)=\mathrm{Cov}\{\xi_{i_1}\xi_{i_2}\}=\sigma_{\zeta}\,\overset{o}{\boldsymbol{\theta}}^{\mathrm{T}}\,\mathbf{S}^{\mathrm{T}}\,\mathbf{I}(i_1-i_2)\mathbf{S}\overset{o}{\boldsymbol{\theta}}\,, \qquad (20)$$

где $\mathbf{I}_p(i_1-i_2)$ – $(p\times p)$-матрица, у которой все элементы равны нулю, кроме элементов одной диагонали, равных единице: если $\Delta=i_1-i_2=0$, то это главная

диагональ; если $\Delta > 0$, то это диагональ, расположенная выше главной диагонали на Δ строк; если $\Delta < 0$, то это диагональ, расположенная ниже главной диагонали на $|\Delta|$ строк.

С учетом (17)–(20) для оценок коэффициентов выполняется

$$\hat{\mathbf{d}} = (\mathbf{R}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}\mathbf{R})^{-1}\mathbf{R}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}\,\mathbf{y}\,. \tag{21}$$

В формулу (17) для матрицы $\mathbf{C}$ входит ненаблюдаемая матрица регрессоров $\mathbf{R}$, а в формулу (18) для матрицы $\boldsymbol{\Sigma}_{\xi}$ – матрица $\boldsymbol{\Psi}$, элементы которой, как следует из (19)–(20), зависят от неизвестных коэффициентов $\overset{o}{\boldsymbol{\theta}}$. Эти обстоятельства используем для построения итерационной процедуры вычисления неизвестных коэффициентов в виде (21).

Пусть матрица $\hat{\mathbf{R}}(r-1)$ – оценка матрицы регрессоров $\mathbf{R}$, полученная на итерации с номером $r-1$; $\hat{\mathbf{d}}(r)$ – оценка вектора коэффициентов $\overset{o}{\boldsymbol{\theta}}$ в виде (21), полученная на итерации с номером r. Тогда для регрессионной модели (15) выполняется

$$\mathbf{y} = \hat{\mathbf{R}}(r-1)\,\hat{\mathbf{d}}(r) + \mathbf{u}(r) = \hat{\mathbf{y}}(r) + \mathbf{u}(r)\,, \tag{22}$$

где $\hat{\mathbf{y}}(r)$ – $(n\times 1)$-вектор выходов, $\mathbf{u}(r)$ – $(n\times 1)$-вектор остатков [15] регрессионной модели.

Итерационная процедура вычисления неизвестных коэффициентов в виде (16)–(17) предусматривает три этапа.

Этап 1. Начальное приближение.

Шаг 1. Формируем матрицы наблюдаемых предыдущих значений переменных (аналогично (11))

$$\mathbf{Z}(p;0) = \begin{bmatrix} x_0 & x_{-1} & \cdots & x_{1-p} \\ x_1 & x_0 & \cdots & x_{2-p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{i-1} & x_{i-2} & \cdots & x_{i-p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n-1} & x_{n-2} & \cdots & x_{n-p} \end{bmatrix}, \tag{23}$$

Шаг 2. Формируем матрицы наблюдаемых регрессоров (аналогично (14))

$$\hat{\mathbf{R}}(0) = \mathbf{Z}(p;0)\,\mathbf{S}\,. \tag{24}$$

Шаг 4. Полагаем $\boldsymbol{\Psi} = \mathbf{O}_{n\times n}$ – нулевая $(n\times n)$-матрица.

Шаг 5. Вычисляем оценку коэффициентов:

$$\hat{\mathbf{d}}(0) = \left([\hat{\mathbf{R}}(0)]^{\mathrm{T}}[\sigma_{\varepsilon}\cdot\mathbf{I}_n + \sigma_{\zeta}\cdot\mathbf{I}_n]^{-1}\,\hat{\mathbf{R}}(0) \right)^{-1} [\hat{\mathbf{R}}(0)]^{\mathrm{T}}[\sigma_{\varepsilon}\cdot\mathbf{I}_n + \sigma_{\zeta}\cdot\mathbf{I}_n]^{-1}\,\mathbf{y}\,. \tag{25}$$

Шаг 6. Вычисляем выходы моделей:

$$\hat{\mathbf{y}}(0) = \hat{\mathbf{R}}(0)\,\hat{\mathbf{d}}(0)\,. \tag{26}$$

Шаг 7. Вычисляем остатки моделей:

$$\mathbf{u}(0)=\mathbf{y}-\hat{\mathbf{y}}(0)\,. \qquad (27)$$

Шаг 8. Вычисляем целевой функционал:

$$\Phi(0)=(\,(n-1)^{-1}\cdot\mathbf{u}^{\mathrm{T}}(0)\mathbf{u}(0)\,)^{1/2}\,. \qquad (28)$$

Этап II. На итерациях $r=1,2,\ldots,r^*$ выполняем такие операции:

Шаг 1. Формируем матрицы оценок предыдущих значений (аналогично (11))

$$\hat{\mathbf{Z}}(p;r-1)=\begin{bmatrix} \hat{y}_0(r-1) & \hat{y}_{-1}(r-1) & \cdots & \hat{y}_{1-p}(r-1) \\ \hat{y}_1(r-1) & \hat{y}_0(r-1) & \cdots & \hat{y}_{2-p}(r-1) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{y}_{i-1}(r-1) & \hat{y}_{i-2}(r-1) & \cdots & \hat{y}_{i-p}(r-1) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{y}_{n-1}(r-1) & \hat{y}_{n-2}(r-1) & \cdots & \hat{y}_{n-p}(r-1) \end{bmatrix}. \qquad (29)$$

Шаг 2. Формируем матрицы регрессоров (аналогично (14))

$$\hat{\mathbf{R}}(r-1)=\mathbf{Z}(p;r-1)\,\mathbf{S}\,. \qquad (30)$$

Шаг 4. Рассчитываем матрицу $\mathbf{\Psi}(r-1)$ – величины $\psi(\Delta;r-1)$, $\Delta=-p+1,-p+2,\ldots\;\ldots,-1,0,1,\ldots,p-2,p-1$ вычисляем по формулам (35)–(36), используя в качестве оценок $\overset{o}{\boldsymbol{\theta}}$ приближения $\hat{\mathbf{d}}(r-1)$, полученные на итерации $r-1$; на итерации $r=1$ используем оценки начального приближения $\hat{\mathbf{d}}(0)$ в (25).

Шаг 5. Вычисляем оценку коэффициентов:

$$\hat{\mathbf{d}}(r)=\left(\;[\hat{\mathbf{R}}\,(r-1)]^{\mathrm{T}}[\sigma_\varepsilon\cdot\mathbf{I}_n+\mathbf{\Psi}(r-1)+\sigma_\zeta\cdot\mathbf{I}_n]^{-1}\,\hat{\mathbf{R}}\,(r-1)\;\right)^{-1}\times$$

$$\times\,[\hat{\mathbf{R}}\,(r-1)]^{\mathrm{T}}[\sigma_\varepsilon\cdot\mathbf{I}_n+\mathbf{\Psi}(r-1)+\sigma_\zeta\cdot\mathbf{I}_n]^{-1}\,\mathbf{y}\,. \qquad (31)$$

Шаг 6. Вычисляем выходы моделей:

$$\hat{\mathbf{y}}(r)=\hat{\mathbf{R}}(r-1)\,\hat{\mathbf{d}}(r)\,. \qquad (32)$$

Шаг 7. Вычисляем остатки моделей:

$$\mathbf{u}(r)=\mathbf{y}-\hat{\mathbf{y}}(r)\,. \qquad (33)$$

Шаг 8. Вычисляем целевой функционал:

$$\Phi(r)=(\,(n-1)^{-1}\cdot\mathbf{u}^{\mathrm{T}}(r)\mathbf{u}(r)\,)^{1/2}\,. \qquad (34)$$

Этап III. Итерационный процесс заканчиваем на итерации r^* при выполнении условия

$$\delta^2=\Phi(r^*-1)-\Phi(r^*)\le\delta_0^2\,, \qquad (35)$$

где δ_0^2 – заданная величина.

Процедура (22)–(35) является частным случаем итерационных процедур [13] и [14], разработанных для оценивания коэффициентов в системах авторегрессионных уравнений: первая из них разработана для случая, когда ковариационные матрицы случайных составляющих в модели функционирования

и модели наблюдения априорно известны, а вторая – для случая, когда они неизвестны.

Для оценки $\hat{\mathbf{d}}$ с учетом (15) и (21) получаем

$$\hat{\mathbf{d}} = [(\mathbf{R}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}\mathbf{R})^{-1}]\mathbf{R}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}\mathbf{R}\overset{o}{\boldsymbol{\theta}} + \mathbf{C}\xi = \overset{o}{\boldsymbol{\theta}} + \mathbf{C}\xi . \quad (36)$$

С учётом (36) регрессионную модель запишем в виде

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{u} = \mathbf{R}\hat{\mathbf{d}} + \mathbf{u} , \quad (37)$$

где $\hat{\mathbf{y}}$ – $(n \times 1)$-вектор выхода регрессионной модели

$$\hat{\mathbf{y}} = \mathbf{R}\hat{\mathbf{d}} = \mathbf{R}\overset{o}{\boldsymbol{\theta}} + \mathbf{R}\mathbf{C}\xi = \overset{o}{\mathbf{y}} + \mathbf{R}\mathbf{C}\xi ; \quad (38)$$

$\mathbf{u}$ – $(n \times 1)$-вектор остатков [15], для которого выполняется

$$\mathbf{u} = \xi - \mathbf{R}\mathbf{C}\xi , \quad E\{\mathbf{u}\} = \mathbf{0}_n , \quad (39)$$

т. е. его математическое ожидание равно нулевому $(n \times 1)$-вектору.

Критерий регулярности МГУА для линейной авторегрессии

Пусть структурная $(p \times \overset{o}{m})$-матрица $\overset{o}{\mathbf{S}}$ соответствует истинной структуре модели объекта (1)–(8), т. е. она однозначно задает структуру авторегрессионного уравнения: указывает, какие именно предыдущие значения определяют текущие значения выходной переменной. Пусть эта структурная матрица неизвестна и её требуется определить по результатам наблюдения функционирования объекта (9), т. е. рассмотрим задачу структурной идентификации. Далее будем предполагать, что для генерации и анализа структур моделей применяется алгоритм полного перебора всех возможных структур моделей, а значение p – число предыдущих значений выходной переменной, которые влияют на её текущее значение – априорно известно.

Пусть структурная $(p \times s)$-матрица $\mathbf{S}$ – набор структурных матриц текущей анализируемой структуры модели.

Рассмотрим так называемый J-функционал качества регрессионного уравнения, отражающий требование минимизации математического ожидания

$$J = E\{(\overset{o}{\mathbf{y}} - \hat{\mathbf{y}})^{\mathrm{T}}(\overset{o}{\mathbf{y}} - \hat{\mathbf{y}})\} , \quad (40)$$

который известен для модели одномерной по выходу регрессии [16, с. 172] и получил в МГУА название “идеальный внешний критерий” [9, с. 99]. Функционал (40) не может применяться при решении практических задач (содержит ненаблюдаемый вектор $\overset{o}{\mathbf{y}}$), но может быть использован для теоретического сравнения методов оценивания, в том числе на основе метода статистических испытаний [12].

Существует ли конструктивная альтернатива ненаблюдаемому вектору $\overset{o}{\mathbf{y}}$ в J-функционале (40), сохраняющая для соответствующего функционала свойства J-функционала? Положительный ответ на этот вопрос для систем статических (одновременных) регрессионных уравнений [12] получен в рамках так называемой схемы повторных наблюдений, которая может быть реализована в условиях активного эксперимента. В этой схеме для заданного вектора значений

входных переменных объекта производится не одно, а пара независимых наблюдения выходных переменных. Первое наблюдение из этой пары участвует в формировании выборки A, другое – выборки B, и, в результате, в схеме повторных наблюдений выполняется $\mathbf{X}(A)=\mathbf{X}(B)$.

Такую схему можно реализовать для статических регрессионных моделей, но для авторегрессионных моделей она принципиально нереализуема: значение каждой из переменных множества X в силу модели (1) формируется с участием случайной ненаблюдаемой составляющей, которую «воспроизвести» нельзя. Поскольку добиться выполнения $\mathbf{X}(A)=\mathbf{X}(B)$ невозможно, то остаётся попытка обеспечить выполнение

$$\mathbf{X}^{\mathrm{T}}(A)\,\mathbf{X}(A)\approx\mathbf{X}^{\mathrm{T}}(B)\,\mathbf{X}(B) \qquad (41)$$

и провести исследование критерия регулярности в этих условиях.

Будем предполагать, что объект, для которого решается задача структурной идентификации, принадлежит к классу объектов, допускающих возможность неоднократного наблюдения реализаций функционирования (например, временные ряды показателей некоторого технологического процесса). И будем выбирать такие реализации, которые начинаются с приблизительно одинаковых начальных условий $\mathbf{x}^{\mathrm{T}}(A,0)\,\mathbf{x}(A,0)\approx\mathbf{x}^{\mathrm{T}}(B,0)\,\mathbf{x}(B,0)$, имеют качественно одинаковый характер переходных процессов и заканчиваются близкими состояниями в конечные моменты времени.

Пусть в качестве двух выборок наблюдений A и B выбраны наблюдения двух реализаций функционирования объекта – две выборки наблюдений выходной переменной, обладающих указанными выше свойствами. Первую выборку A будем называть обучающей, а вторую B – проверочной. На обучающей выборке будем оценивать коэффициенты в авторегрессионном уравнении с текущей анализируемой структурой, а на проверочной будем оценивать качество построенной модели. (В дальнейшем будем называть такой способ формирования обучающей и проверочной выборок «схемой квазиповторных наблюдений».).

В соответствии с (39) для $(n(B)\times 1)$-вектора остатков на выборке B выполняется

$$\mathbf{u}(B/A,S)=\mathbf{y}(B)-\hat{\mathbf{y}}(B\,|\,A,S)=\mathbf{y}(B)-\mathbf{R}(B,S)\,\hat{\mathbf{d}}(A,S), \qquad (42)$$

где $\mathbf{y}(B)$ – $(n(B)\times 1)$-вектор наблюдений выходной переменной выборки B; $\hat{\mathbf{y}}(B\,|\,A,S)$ – $(n(B)\times 1)$-вектор выходов регрессионной модели на выборке B, рассчитанный по модели, оценки коэффициентов которой $\hat{\mathbf{d}}(A,S)$ получены в соответствии с (21)–(41) на обучающей выборке A для структуры S; $n(B)$ – объем проверочной выборки.

В соответствии с (37)–(39) для $(n(B)\times 1)$-вектора остатков $\mathbf{u}(B\,|\,A,S)$ выполняется

$$\mathbf{u}(B\,|\,A,S)=\overset{o}{\mathbf{y}}(B)+\boldsymbol{\xi}(B)-\mathbf{R}(B,S)\,\mathbf{C}(A,S)\,\overset{o}{\mathbf{y}}(A)-\mathbf{R}(B,S)\,\mathbf{C}(A,S)\,\boldsymbol{\xi}(A)=$$
$$=\boldsymbol{\delta}(B\,|\,A,S)+\boldsymbol{\xi}(B)-\mathbf{R}(B,S)\,\mathbf{C}(A,S)\,\boldsymbol{\xi}(A), \qquad (43)$$

где

$$\boldsymbol{\delta}(B\,|\,A,S)=\overset{\circ}{\mathbf{y}}(B)-\mathbf{R}(B,S)\,\mathbf{C}(A,S)\,\overset{\circ}{\mathbf{y}}(A) \tag{44}$$

– так называемый $(n\times 1)$-вектор смещения, обусловленный выбором структуры S вместо $\overset{\circ}{S}$.

Определение 1. Случайная величина

$$ARD^*(S)=\mathbf{u}^{\mathrm{T}}(B\,|\,A,S)\mathbf{u}(B\,|\,A,S) \tag{45}$$

называется критерием регулярности для авторегрессионного уравнения, где звёздочка означает тот факт, что выборки A и B получены в условиях квазиповторных наблюдений.

Определение 2. Оптимальным множеством регрессоров называется множество регрессоров, соответствующее набору структурных матриц S_0:

$$S_0=\arg\min_{S\subseteq S^*(p)} E\{ARD^*(S)\}, \tag{46}$$

где $S^*(p)$ – множество возможных структурных матриц при заданном параметре p.

Определение 3. Оптимальным по количеству и составу регрессоров называется авторегрессионное уравнение, построенное на множестве регрессоров, соответствующем структурной матрице S_0.

Для математического ожидания $ARD^*(S)$, с учётом результатов [12], выполняется

$$\begin{aligned}
&E\{ARD^*(S)\}=E\{\mathbf{u}^{\mathrm{T}}(B\,|\,A,S)\mathbf{u}(B\,|\,A,S)\}=\\
&=\boldsymbol{\delta}^{\mathrm{T}}(B\,|\,A,S)\,\boldsymbol{\delta}(B\,|\,A,S)+E\,\{[\boldsymbol{\xi}^{\mathrm{T}}(B)\boldsymbol{\xi}(B)]\}-\\
&-E\{\boldsymbol{\xi}^{\mathrm{T}}(B)\,\mathbf{R}(B,S)\,\mathbf{C}(A,S)\,\boldsymbol{\xi}(A)\}-E\{[\mathbf{R}(B,S)\,\mathbf{C}(A,S)\,\boldsymbol{\xi}(A)]^{\mathrm{T}}\boldsymbol{\xi}(B)\}+\\
&\qquad+E\{[\mathbf{R}(B,S)\,\mathbf{C}(A,S)\boldsymbol{\xi}(A)]^{\mathrm{T}}\mathbf{R}(B,S)\,\mathbf{C}(A,S)\,\boldsymbol{\xi}(A)\}.
\end{aligned} \tag{47}$$

Для второго слагаемого в (47), с учётом (18)–(20), выполняется

$$E\,\{\boldsymbol{\xi}^{\mathrm{T}}(B)\boldsymbol{\xi}(B)\}=n(B)\cdot(\sigma_\varepsilon+\psi(0,S)+\sigma_\zeta\,), \tag{48}$$

а третье и четвертое слагаемые равны нулю, поскольку $\boldsymbol{\xi}(A)$ и $\boldsymbol{\xi}(B)$ независимы.

Для пятого слагаемого в (47), учитывая (17), получаем

$$\begin{aligned}
&E\{[\mathbf{R}(B,S)\,\mathbf{C}(A,S)\boldsymbol{\xi}(A)]^{\mathrm{T}}\mathbf{R}(B,S)\,\mathbf{C}(A,S)\,\boldsymbol{\xi}(A)\}=\\
&\qquad=\mathrm{tr}\,[\,\mathbf{R}(B,S)\,(\mathbf{R}^{\mathrm{T}}(A,S)\,\boldsymbol{\Sigma}_\xi^{-1}\mathbf{R}(A,S))^{-1}\,\mathbf{R}^{\mathrm{T}}(B,S)\,].
\end{aligned} \tag{49}$$

Подставляя в (47) выражения (48) и (49), получаем

$$\begin{aligned}
&E\{ARD^*(S)\}=\boldsymbol{\delta}^{\mathrm{T}}(B\,|\,A,S)\,\boldsymbol{\delta}(B\,|\,A,S)+n(B)\cdot(\sigma_\varepsilon+\psi_{kq}(0)+\sigma_\zeta)+\\
&+\mathrm{tr}\,[\,\mathbf{R}(B,S)\,(\mathbf{R}^{\mathrm{T}}(A,S)\,\boldsymbol{\Sigma}_\xi^{-1}\mathbf{R}(A,S))^{-1}\,\mathbf{R}^{\mathrm{T}}(B,S)\,]=\\
&\qquad=\Delta(B\,|\,A,S)+n(B)\cdot(\sigma_\varepsilon+\psi(0,S)+\sigma_\zeta)+\mathrm{tr}\,[\,\mathbf{P}(B\,|\,A,S)\,],
\end{aligned} \tag{50}$$

где $\Delta(B\,|\,A,S)$ – смещение, обусловленное выбором структуры S вместо $\overset{\circ}{S}$:

$$\Delta(B\,|\,A,S)=\boldsymbol{\delta}^{\mathrm{T}}(B\,|\,A,S)\,\boldsymbol{\delta}(B\,|\,A,S)\,; \tag{51}$$

$$\mathbf{P}(B\,|\,A,S)=\mathbf{R}(B,S)\,(\mathbf{R}^{\mathrm{T}}(A,S)\,\boldsymbol{\Sigma}_\xi^{-1}\mathbf{R}(A,S))^{-1}\,\mathbf{R}^{\mathrm{T}}(B,S)\,. \tag{52}$$

Исследование критерия регулярности МГУА

Установим свойства критерия регулярности МГУА (45). С этой целью исследуем, как изменяется математическое ожидание критерия в зависимости от состава множества регрессоров. В случае истинной структуры получаем

$$E\{ARD^*(\overset{o}{S})\} = n(B)\cdot(\sigma_\varepsilon + \psi_{kq}(0) + \sigma_\zeta) +$$
$$+ \operatorname{tr}[\mathbf{R}(B,\overset{o}{S})(\mathbf{R}^{\mathrm{T}}(A,\overset{o}{S})\boldsymbol{\Sigma}_\xi^{-1}\mathbf{R}(A,\overset{o}{S}))^{-1}\mathbf{R}^{\mathrm{T}}(B,\overset{o}{S})] =$$
$$= n(B)\cdot(\sigma_\varepsilon + \psi(0,\overset{o}{S}) + \sigma_\zeta) + \operatorname{tr}[\mathbf{P}(B\,|\,A,\overset{o}{S})]. \quad (53)$$

Случай недостающего регрессора. Рассмотрим случай, когда в текущей структуре пропущен один регрессор. Предположим для простоты, что это регрессор с номером p, он является максимально удаленным предыдущим значением переменной, участвующим в формировании её текущего значения, т. е. для структурных матриц выполняется

$$\overset{o}{\mathbf{S}} = [\mathbf{S} \,\vdots\, \mathbf{s}], \quad (54)$$

где $\overset{o}{\mathbf{S}}$ – структурная $(p\times\overset{o}{m})$-матрица истинной модели; $\mathbf{S}$ – структурная $(p\times(\overset{o}{m}-1))$-матрица текущей модели; $\mathbf{s}$ – $(p\times 1)$-вектор, для которого выполняется

$$\mathbf{s} = (0, 0, ..., 0, 1)^{\mathrm{T}}. \quad (55)$$

Другими словами, в модели функционирования (4) в формировании величины $\overset{*}{x}_i$ участвует величина $\overset{*}{x}_{i-p}$, но в текущую модель она не включена.

Для матриц регрессоров, соответствующих наборам $\overset{o}{S}$ и S в (54), выполняется

$$\mathbf{R}(\overset{o}{S}) = \mathbf{X}\overset{o}{\mathbf{S}} = \mathbf{X}[\mathbf{S} \,\vdots\, \mathbf{s}] = [\mathbf{X}\mathbf{S} \,\vdots\, \mathbf{X}\mathbf{s}] = [\mathbf{R}(S) \,\vdots\, \mathbf{m}], \quad (56)$$

где $\mathbf{m}$ – $(n\times 1)$-вектор наблюдений пропущенного регрессора.

Используя (50)–(53), вычислим разность математических ожиданий критерия регулярности для текущей структуры S и истинной структуры $\overset{o}{S}$:

$$\Delta_1(S,\overset{o}{S}) = E\{ARD^*(S)\} - E\{ARD^*(\overset{o}{S})\} = \Delta(B\,|\,A,S) + \operatorname{tr}[\mathbf{P}(B\,|\,A,S)] - \operatorname{tr}[\mathbf{P}(B\,|\,A,\overset{o}{S})],$$
(57)

где $\Delta(B\,|\,A,S) = \boldsymbol{\delta}^{\mathrm{T}}(B\,|\,A,S)\boldsymbol{\delta}(B\,|\,A,S)$ – скалярная величина; $\boldsymbol{\delta}(B\,|\,A,S)$ – введенный в (44) $(n\times 1)$-вектор смещения, обусловленный выбором структуры S вместо $\overset{o}{S}$.

Для вектора смещения $\boldsymbol{\delta}(B\,|\,A,S)$ выполняется

$$\boldsymbol{\delta}(B\,|\,A,S) = \overset{o}{\mathbf{y}}(B) - \mathbf{R}(B,S)\mathbf{C}(A,S)\overset{o}{\mathbf{y}}(A) = \mathbf{R}(B,\overset{o}{S})\overset{o}{\boldsymbol{\theta}} - \mathbf{R}(B,S)\mathbf{C}(A,S)\mathbf{R}(A,\overset{o}{S})\overset{o}{\boldsymbol{\theta}} =$$
$$= \mathbf{R}(B,\overset{o}{S})\overset{o}{\boldsymbol{\theta}} - \mathbf{R}(B,S)(\mathbf{R}^{\mathrm{T}}(A,S)\boldsymbol{\Sigma}_\xi^{-1}\mathbf{R}(A,S))^{-1}\mathbf{R}^{\mathrm{T}}(A,S)\boldsymbol{\Sigma}_\xi^{-1}\mathbf{R}(A,\overset{o}{S})\overset{o}{\boldsymbol{\theta}}. \quad (58)$$

Запишем (58) с учётом (56)

$$\boldsymbol{\delta}(B\,|\,A,S)=\left[\mathbf{R}(B,S)\ \vdots\ \mathbf{m}(B)\right]\overset{\circ}{\boldsymbol{\theta}}-$$

$$-\mathbf{R}(B,S)(\mathbf{R}^{\mathrm{T}}(A,S)\boldsymbol{\Sigma}_{\xi}^{-1}\mathbf{R}(A,S))^{-1}\mathbf{R}^{\mathrm{T}}(A,S)\boldsymbol{\Sigma}_{\xi}^{-1}\left[\mathbf{R}(A,S)\ \vdots\ \mathbf{m}(A)\right]\overset{\circ}{\boldsymbol{\theta}}=$$

$$=\left[\mathbf{R}(B,S)\ \vdots\ \mathbf{m}(B)\right]\overset{\circ}{\boldsymbol{\theta}}-$$

$$\left[\mathbf{R}(B,S)\ \vdots\ \mathbf{R}(B,S)(\mathbf{R}^{\mathrm{T}}(A,S)\boldsymbol{\Sigma}_{\xi}^{-1}\mathbf{R}(A,S))^{-1}\mathbf{R}^{\mathrm{T}}(A,S)\boldsymbol{\Sigma}_{\xi}^{-1}\mathbf{m}(A)\right]\overset{\circ}{\boldsymbol{\theta}}=$$

$$=[\,\mathbf{O}_{n\times(\overset{\circ}{m}-1)}\ \vdots\ \mathbf{m}(B)-\mathbf{P}(B,A,S)\,\boldsymbol{\Sigma}_{\xi}^{-1}\,\mathbf{m}(A)]\,\overset{\circ}{\boldsymbol{\theta}},\tag{59}$$

$$\mathbf{P}(B,A,S)=\mathbf{R}(B,S)\,(\mathbf{R}^{\mathrm{T}}(A,S)\boldsymbol{\Sigma}_{\xi}^{-1}\mathbf{R}(A,S))^{-1}\mathbf{R}^{\mathrm{T}}(A,S).\tag{60}$$

Учитывая (59) и (60) для величины $\Delta(B\,|\,A,S)$, введённой в (57), получаем

$$\Delta(B\,|\,A,S)=\overset{\circ}{\boldsymbol{\theta}}^{\mathrm{T}}\begin{bmatrix}\mathbf{O}_{(\overset{\circ}{m}-1)\times(\overset{\circ}{m}-1)} & \mathbf{0}_{\overset{\circ}{m}-1}\\ \mathbf{0}^{\mathrm{T}}_{\overset{\circ}{m}-1} & \begin{array}{l}(\,\mathbf{m}^{\mathrm{T}}(B)-\mathbf{m}^{\mathrm{T}}(A)\ \boldsymbol{\Sigma}_{\xi}^{-1}\,\mathbf{P}^{\mathrm{T}}(B\,|\,A,S))\times\\ \times(\,\mathbf{m}(B)-\mathbf{P}(B\,|\,A,S)\,\boldsymbol{\Sigma}_{\xi}^{-1}\,\mathbf{m}(A))\end{array}\end{bmatrix}\overset{\circ}{\boldsymbol{\theta}}.\tag{61}$$

Для $(\overset{\circ}{m},\overset{\circ}{m})$-го элемента матрицы в (61) выполняется

$$a=(\,\mathbf{m}^{\mathrm{T}}(B)-\mathbf{m}^{\mathrm{T}}(A)\ \boldsymbol{\Sigma}_{\xi}^{-1}\,\mathbf{P}^{\mathrm{T}}(B\,|\,A,S))\ (\,\mathbf{m}(B)-\mathbf{P}(B\,|\,A,S)\,\boldsymbol{\Sigma}_{\xi}^{-1}\,\mathbf{m}(A))=$$

$$=\mathbf{m}^{\mathrm{T}}(B)\,\mathbf{m}(B)-\mathbf{m}^{\mathrm{T}}(B)\,\mathbf{R}(B,S)\,(\mathbf{R}^{\mathrm{T}}(A,S)\,\boldsymbol{\Sigma}_{\xi}^{-1}\mathbf{R}(A,S))^{-1}\,\mathbf{R}^{\mathrm{T}}(A,S)\,\boldsymbol{\Sigma}_{\xi}^{-1}\,\mathbf{m}(A)-$$

$$-\mathbf{m}^{\mathrm{T}}(A)\,\boldsymbol{\Sigma}_{\xi}^{-1}\ \mathbf{R}(A,S)\,(\mathbf{R}^{\mathrm{T}}(A,S)\,\boldsymbol{\Sigma}_{\xi}^{-1}\mathbf{R}(A,S))^{-1}\,\mathbf{R}^{\mathrm{T}}(B,S)\,\mathbf{m}(B)+$$

$$+\mathbf{m}^{\mathrm{T}}(A)\ \boldsymbol{\Sigma}_{\xi}^{-1}\mathbf{R}(A,S)\,(\mathbf{R}^{\mathrm{T}}(A,S)\,\boldsymbol{\Sigma}_{\xi}^{-1}\mathbf{R}(A,S))^{-1}\mathbf{R}^{\mathrm{T}}(B,S)\times$$

$$\times\mathbf{R}(B,S)\,(\mathbf{R}^{\mathrm{T}}(A,S)\,\boldsymbol{\Sigma}_{\xi}^{-1}\mathbf{R}(A,S))^{-1}\,\mathbf{R}^{\mathrm{T}}(A,S)\,\boldsymbol{\Sigma}_{\xi}^{-1}\,\mathbf{m}(A).\tag{62}$$

Далее используем то обстоятельство, что обучающая (A) и проверочная (B) выборки получены особым способом как пара реализаций функционирования объекта с близкими начальными условиями, качественно одинаковым характером переходных процессов и близкими состояниями в конечные моменты времени.

Введём в рассмотрение разности:

$$\mathbf{G}(X)=\mathbf{X}^{\mathrm{T}}(B)\ \mathbf{X}(B)-\mathbf{X}^{\mathrm{T}}(A)\ \mathbf{X}(A),\tag{63}$$

$$g(\mathbf{m})=\mathbf{m}^{\mathrm{T}}(B)\,\mathbf{m}(B)-\mathbf{m}^{\mathrm{T}}(A)\,\mathbf{m}(A),\tag{64}$$

$$\mathbf{G}(R,S)=\mathbf{R}^{\mathrm{T}}(B,S)\ \mathbf{R}(B,S)-\mathbf{R}^{\mathrm{T}}(A,S)\ \mathbf{R}(A,S),\tag{65}$$

$$\mathbf{g}^{\mathrm{T}}(R,S)=\mathbf{m}^{\mathrm{T}}(B)\ \mathbf{R}(B,S)-\mathbf{m}^{\mathrm{T}}(A)\ \mathbf{R}(A,S).\tag{66}$$

(Влияние скаляра $g(\mathbf{m})$, вектора $\mathbf{g}(R,S)$, матрицы $\mathbf{G}(R,S)$ на условие упрощения структуры по числу регрессоров оценены методом статистических испытаний в отдельном исследовании.).

Учитывая (63)–(66), для (62) получаем

$$a=\mathbf{m}^{\mathrm{T}}(A)\,[\,\mathbf{M}^{\mathrm{T}}(A,S)\,\mathbf{M}(A,S)\,]\,\mathbf{m}(A)+\rho_1(\mathbf{m}),\tag{67}$$

где $\mathbf{M}(A,S)$ – $(n(A)\times n(A))$-матрица

$$\mathbf{M}(A,S)=[\,\mathbf{I}_{n(A)}-\mathbf{R}(A,S)(\mathbf{R}^{\mathrm{T}}(A,S)\,\boldsymbol{\Sigma}_{\xi}^{-1}\,\mathbf{R}(A,S))^{-1}\mathbf{R}^{\mathrm{T}}(A,S)\,\boldsymbol{\Sigma}_{\xi}^{-1}\,],\tag{68}$$

а для скаляра $\rho_1(\mathbf{m})$ выполняется

$$\rho_1(\mathbf{m}) = g(\mathbf{m}) - \mathbf{g}^{\mathrm{T}}(R,S)(\mathbf{R}^{\mathrm{T}}(A,S)\,\boldsymbol{\Sigma}_\xi^{-1}\mathbf{R}(A,S))^{-1}\,\mathbf{R}^{\mathrm{T}}(A,S)\,\boldsymbol{\Sigma}_\xi^{-1}\,\mathbf{m}(A) -$$

$$- \mathbf{m}^{\mathrm{T}}(A)\ \boldsymbol{\Sigma}_\xi^{-1}\ \mathbf{R}(A,S)(\mathbf{R}^{\mathrm{T}}(A,S)\,\boldsymbol{\Sigma}_\xi^{-1}\mathbf{R}(A,S))^{-1}\,\mathbf{g}(R,S) +$$

$$+ \mathbf{m}^{\mathrm{T}}(A)\ \boldsymbol{\Sigma}_\xi^{-1}\mathbf{R}(A,S)(\mathbf{R}^{\mathrm{T}}(A,S)\,\boldsymbol{\Sigma}_\xi^{-1}\mathbf{R}(A,S))^{-1}\mathbf{G}(R,S) \times$$

$$\times (\mathbf{R}^{\mathrm{T}}(A,S)\,\boldsymbol{\Sigma}_\xi^{-1}\mathbf{R}(A,S))^{-1}\,\mathbf{R}^{\mathrm{T}}(A,S)\,\boldsymbol{\Sigma}_\xi^{-1}\,\mathbf{m}(A) =$$

$$= g(\mathbf{m}) - \mathbf{g}^{\mathrm{T}}(R,S)\,\mathbf{C}(A,S)\,\mathbf{m}(A) - \mathbf{m}^{\mathrm{T}}(A)\,\mathbf{C}^{\mathrm{T}}(A,S)\,\mathbf{g}(R,S) +$$

$$+ \mathbf{m}^{\mathrm{T}}(A)\,\mathbf{C}^{\mathrm{T}}(A,S)\,\mathbf{G}(R,S)\,\mathbf{C}(A,S)\,\mathbf{m}(A). \qquad (69)$$

Учитывая (54), (67), (68) и соотношение

$$\mathring{\boldsymbol{\theta}}^{\mathrm{T}} = (\mathring{\theta}_1, \mathring{\theta}_2, \ldots, \mathring{\theta}_{\mathring{m}}), \qquad (70)$$

получаем

$$\Delta = \mathring{\boldsymbol{\theta}}^{\mathrm{T}} \left[\begin{array}{c|c} \mathbf{O}_{(\mathring{m}-1)\times(\mathring{m}-1)} & \mathbf{0}_{\mathring{m}-1} \\ \hline \mathbf{0}^{\mathrm{T}}_{\mathring{m}-1} & \mathbf{m}^{\mathrm{T}}(A)\,\mathbf{M}^{\mathrm{T}}(A,S)\,\mathbf{M}(A,S)\,\mathbf{m}(A) + \rho_1(\mathbf{m}) \end{array} \right] \mathring{\boldsymbol{\theta}} =$$

$$= (\mathring{\theta}_{\mathring{m}})^2 \cdot \mathbf{m}^{\mathrm{T}}(A)\,\mathbf{H}(A,S)\,\mathbf{m}(A) + (\mathring{\theta}_{\mathring{m}})^2 \cdot \rho_1(\mathbf{m}), \qquad (71)$$

$$\mathbf{H}(A,S) = \mathbf{M}^{\mathrm{T}}(A,S)\,\mathbf{M}(A,S). \qquad (72)$$

Итак, в (59)–(72) установлено

$$\Delta(B \mid A,S) = \boldsymbol{\delta}^{\mathrm{T}}(B \mid A,S)\,\boldsymbol{\delta}(B \mid A,S) = (\mathring{\theta}_{\mathring{m}})^2 \cdot \mathbf{m}^{\mathrm{T}}(A)\,\mathbf{H}(A,S)\,\mathbf{m}(A) + (\mathring{\theta}_{\mathring{m}})^2 \cdot \rho_1(\mathbf{m}). \qquad (73)$$

Используя (73), для разности (57) получаем

$$\Delta_1(S,\mathring{S}) = \Delta(B \mid A,S) + \mathrm{tr}[\,\mathbf{P}(B \mid A,S)] - \mathrm{tr}[\,\mathbf{P}(B \mid A,\mathring{S})]. \qquad (74)$$

Теперь, продолжая вычисление $\Delta_1(S,\mathring{S})$, установим разность следов в (74). Учитывая (56), получаем

$$\mathbf{P}(B \mid A,S) - \mathbf{P}(B \mid A,\mathring{S}) =$$

$$= \mathbf{R}(B,S)(\mathbf{R}^{\mathrm{T}}(A,S)\,\boldsymbol{\Sigma}_\xi^{-1}\mathbf{R}(A,S))^{-1}\,\mathbf{R}^{\mathrm{T}}(B,S) -$$

$$\mathbf{R}(B,\mathring{S})(\mathbf{R}^{\mathrm{T}}(A,\mathring{S})\,\boldsymbol{\Sigma}_\xi^{-1}\mathbf{R}(A,\mathring{S}))^{-1}\,\mathbf{R}^{\mathrm{T}}(B,\mathring{S}) =$$

$$= \mathbf{R}(B,S)[\,\mathbf{R}^{\mathrm{T}}(A,S)\,\boldsymbol{\Sigma}_\xi^{-1}\mathbf{R}(A,S)\,]^{-1}\mathbf{R}^{\mathrm{T}}(B,S) -$$

$$- \left[\mathbf{R}(B,S) \,\vdots\, \mathbf{m}(B)\right] \left[\begin{array}{c|c} \mathbf{R}^{\mathrm{T}}(A,S)\,\boldsymbol{\Sigma}_\xi^{-1}\,\mathbf{R}(A,S) & \mathbf{R}^{\mathrm{T}}(A,S)\,\boldsymbol{\Sigma}_\xi^{-1}\,\mathbf{m}(A) \\ \hline \mathbf{m}^{\mathrm{T}}(A)\,\boldsymbol{\Sigma}_\xi^{-1}\,\mathbf{R}(A,S) & \mathbf{m}^{\mathrm{T}}(A)\,\boldsymbol{\Sigma}_\xi^{-1}\,\mathbf{m}(A) \end{array} \right]^{-1} \left[\begin{array}{c} \mathbf{R}^{\mathrm{T}}(B,S) \\ \hline \mathbf{m}^{\mathrm{T}}(B) \end{array} \right]. \qquad (75)$$

Для перемножения блочных матриц в (75) применим формулу обращения блочной матрицы (она является частным случаем формулы Фробениуса [16, с. 302]):

$$\mathbf{A}^{-1} = \left[\begin{array}{c|c} \mathbf{B} & \mathbf{c} \\ \hline \mathbf{d}^{\mathrm{T}} & e \end{array} \right]^{-1} = \left[\begin{array}{c|c} \mathbf{B}^{-1} + \mathbf{B}^{-1}\mathbf{c} \cdot f^{-1} \cdot \mathbf{d}^{\mathrm{T}}\mathbf{B}^{-1} & -\mathbf{B}^{-1}\mathbf{c} \cdot f^{-1} \\ \hline -f^{-1} \cdot \mathbf{d}^{\mathrm{T}}\mathbf{B}^{-1} & f^{-1} \end{array} \right], \quad f = e - \mathbf{d}^{\mathrm{T}}\mathbf{B}^{-1}\mathbf{c}. \qquad (76)$$

В нашем случае выполняется

$$\mathbf{B} = \mathbf{R}^{\mathrm{T}}(A,S)\,\boldsymbol{\Sigma}_\xi^{-1}\,\mathbf{R}(A,S), \quad \mathbf{c} = \mathbf{R}^{\mathrm{T}}(A,S)\,\boldsymbol{\Sigma}_\xi^{-1}\,\mathbf{m}(A), \qquad (77)$$

$$\mathbf{d}^{\mathrm{T}} = \mathbf{m}^{\mathrm{T}}(A)\,\mathbf{\Sigma}_{\xi}^{-1}\,\mathbf{R}(A,S),\quad e = \mathbf{m}^{\mathrm{T}}(A)\,\mathbf{\Sigma}_{\xi}^{-1}\,\mathbf{m}(A), \tag{78}$$

$$f = \mathbf{m}^{\mathrm{T}}(A)\,\mathbf{\Sigma}_{\xi}^{-1}\,\mathbf{m}(A) - \mathbf{m}^{\mathrm{T}}(A)\,\mathbf{\Sigma}_{\xi}^{-1}\,\mathbf{R}(A,S)\,(\mathbf{R}^{\mathrm{T}}(A,S)\,\mathbf{\Sigma}_{\xi}^{-1}\,\mathbf{R}(A,S))^{-1}\,\mathbf{R}^{\mathrm{T}}(A,S)\,\mathbf{\Sigma}_{\xi}^{-1}\,\mathbf{m}(A) =$$
$$= \mathbf{m}^{\mathrm{T}}(A)\,\mathbf{\Sigma}_{\xi}^{-1}\,\mathbf{M}(A,S)\,\mathbf{m}(A). \tag{79}$$

Продолжая (75) и учитывая (76)–(79), получаем

$$\mathbf{P}(B,A,S) - \mathbf{P}(B,A,\overset{o}{S}) = \left[\mathbf{R}(B,S) \;\vdots\; \mathbf{m}(B)\right]\times$$
$$\times\begin{bmatrix} \mathbf{B}^{-1} + \\ +\mathbf{B}^{-1}\mathbf{R}^{\mathrm{T}}(A,S)\mathbf{\Sigma}_{\xi}^{-1}\mathbf{m}(A)\cdot f^{-1}\cdot\mathbf{m}^{\mathrm{T}}(A)\mathbf{\Sigma}_{\xi}^{-1}\mathbf{R}(A,S)\,\mathbf{B}^{-1} & -\mathbf{B}^{-1}\mathbf{R}^{\mathrm{T}}(A,S)\mathbf{\Sigma}_{\xi}^{-1}\mathbf{m}(A)\cdot f^{-1} \\ -f^{-1}\cdot\mathbf{m}^{\mathrm{T}}(A)\,\mathbf{\Sigma}_{\xi}^{-1}\,\mathbf{R}(A,S)\,\mathbf{B}^{-1} & f^{-1} \end{bmatrix}\begin{bmatrix}\mathbf{R}^{\mathrm{T}}(B,S) \\ \mathbf{m}^{\mathrm{T}}(B)\end{bmatrix} =$$
$$= -f^{-1}\cdot\mathbf{R}(B,S)\,\mathbf{B}^{-1}\mathbf{R}^{\mathrm{T}}(A,S)\mathbf{\Sigma}_{\xi}^{-1}\mathbf{m}(A)\,\mathbf{m}^{\mathrm{T}}(A)\,\mathbf{\Sigma}_{\xi}^{-1}\mathbf{R}(A,S)\,\mathbf{B}^{-1}\mathbf{R}^{\mathrm{T}}(B,S) +$$
$$+ f^{-1}\cdot\mathbf{R}(B,S)\,\mathbf{B}^{-1}\mathbf{R}^{\mathrm{T}}(A,S)\mathbf{\Sigma}_{\xi}^{-1}\boldsymbol{\delta}_{R}(A)\,\boldsymbol{\delta}_{R}^{\mathrm{T}}(B) +$$
$$+ f^{-1}\cdot\mathbf{m}(B)\,\mathbf{m}^{\mathrm{T}}(A)\,\mathbf{\Sigma}_{\xi}^{-1}\,\mathbf{R}(A,S)\,\mathbf{B}^{-1}\mathbf{R}^{\mathrm{T}}(B,S) - f^{-1}\cdot\mathbf{m}(B)\,\mathbf{m}^{\mathrm{T}}(B). \tag{80}$$

Учитывая (80), для (75) получаем

$$\mathrm{tr}\,[\,\mathbf{P}(B\,|\,A,S) - \mathbf{P}(B\,|\,\overset{o}{A},S)] = -f^{-1}\cdot g(\mathbf{m}) - f^{-1}\cdot\mathbf{m}^{\mathrm{T}}(A)\,\mathbf{m}(A) +$$
$$+ f^{-1}\cdot\mathbf{g}^{\mathrm{T}}(R,S)\,\mathbf{B}^{-1}\mathbf{R}^{\mathrm{T}}(A,S)\mathbf{\Sigma}_{\xi}^{-1}\,\mathbf{m}(A) + f^{-1}\cdot\mathbf{m}^{\mathrm{T}}(A)\,\mathbf{R}(A,S)\,\mathbf{B}^{-1}\mathbf{R}^{\mathrm{T}}(A,S)\mathbf{\Sigma}_{\xi}^{-1}\,\mathbf{m}(A) +$$
$$+ f^{-1}\cdot\mathbf{m}^{\mathrm{T}}(A)\,\mathbf{\Sigma}_{\xi}^{-1}\mathbf{R}(A,S)\,\mathbf{B}^{-1}\mathbf{g}(R,S) + f^{-1}\cdot\mathbf{m}^{\mathrm{T}}(A)\,\mathbf{\Sigma}_{\xi}^{-1}\mathbf{R}(A,S)\,\mathbf{B}^{-1}\mathbf{R}^{\mathrm{T}}(A,S)\mathbf{m}(A) -$$
$$- f^{-1}\cdot\mathbf{m}^{\mathrm{T}}(A)\,\mathbf{\Sigma}_{\xi}^{-1}\underline{\underline{\mathbf{R}}}(A,S)\,\mathbf{B}^{-1}\mathbf{G}(R,S)\,\mathbf{B}^{-1}\underline{\underline{\mathbf{R}}}^{\mathrm{T}}(A,S)\mathbf{\Sigma}_{\xi}^{-1}\,\mathbf{m}(A) -$$
$$- f^{-1}\cdot\mathbf{m}^{\mathrm{T}}(A)\,\mathbf{\Sigma}_{\xi}^{-1}\mathbf{R}(A,S)\,\mathbf{B}^{-1}\,\mathbf{R}^{\mathrm{T}}(A,S)\,\mathbf{R}(A,S)\,\mathbf{B}^{-1}\,\mathbf{R}^{\mathrm{T}}(A,S)\mathbf{\Sigma}_{\xi}^{-1}\,\mathbf{m}(A) =$$
$$= -f^{-1}\cdot\mathbf{m}^{\mathrm{T}}(A)\,\mathbf{H}(A,S)\,\mathbf{m}(A) - f^{-1}\cdot q(\mathbf{m}), \tag{81}$$

где $\mathbf{H}(A,S) = \mathbf{M}^{\mathrm{T}}(A,S)\,\mathbf{M}(A,S)$; $\mathbf{M}(A,S)$ – $(n(A)\times n(A))$-матрица введена в (67)–(68); а для скалярной величины $q(\mathbf{m})$ выполняется

$$q(\mathbf{m}) = g(\mathbf{m}) - \mathbf{g}^{\mathrm{T}}(R)\,\mathbf{B}^{-1}\mathbf{R}^{\mathrm{T}}(A,S)\mathbf{\Sigma}_{\xi}^{-1}\,\mathbf{m}(A) - \mathbf{m}^{\mathrm{T}}(A)\,\mathbf{\Sigma}_{\xi}^{-1}\mathbf{R}(A,S)\,\mathbf{B}^{-1}\mathbf{g}(R) +$$
$$+ \mathbf{m}^{\mathrm{T}}(A)\,\mathbf{\Sigma}_{\xi}^{-1}\mathbf{R}(A,S)\,\mathbf{B}^{-1}\mathbf{G}(R)\,\mathbf{B}^{-1}\mathbf{R}^{\mathrm{T}}(A,S)\mathbf{\Sigma}_{\xi}^{-1}\,\mathbf{m}(A) =$$
$$= g(\mathbf{m}) - \mathbf{g}^{\mathrm{T}}(R)\,\mathbf{C}(A,S)\,\mathbf{m}(A) - \mathbf{m}^{\mathrm{T}}(A)\,\mathbf{C}^{\mathrm{T}}(A,S)\,\mathbf{g}^{\mathrm{T}}(R) +$$
$$+ \mathbf{m}^{\mathrm{T}}(A)\,\mathbf{C}^{\mathrm{T}}(A,S)\,\mathbf{G}(R)\,\mathbf{C}(A,S)\,\mathbf{m}(A). \tag{82}$$

Подставляя (81) в (74), получаем

$$\Delta_1(S,\overset{o}{S}) = (\overset{o}{\theta}{}_{\overset{o}{m}})^2\cdot\mathbf{m}^{\mathrm{T}}(A)\,\mathbf{H}(A,S)\,\mathbf{m}(A) + (\overset{o}{\theta}{}_{\overset{o}{m}})^2\cdot\rho_1(\mathbf{m}) -$$
$$- f^{-1}\cdot\mathbf{m}^{\mathrm{T}}(A)\,\mathbf{H}(A,S)\,\mathbf{m}(A) - f^{-1}\cdot q(\mathbf{m}). \tag{83}$$

Если $\Delta_1(S,\overset{o}{S}) > 0$, то структура $\overset{o}{S}$ лучше S; если $\Delta_1(S,\overset{o}{S}) < 0$, то структура S лучше $\overset{o}{S}$; если $\Delta_1(S,\overset{o}{S}) = 0$, то структура S лучше $\overset{o}{S}$ по дополнительному принципу простоты. Выполнение $\Delta_1(S,\overset{o}{S}) \le 0$ является условием так называемой редукции (упрощения) модели, оптимальной по структуре. Из (83) для условия редукции получаем

$$(\overset{o}{\theta}{}^{o}_{m})^{2}\cdot\mathbf{m}^{\mathrm{T}}(A)\,\mathbf{H}(A,S)\,\mathbf{m}(A)+(\overset{o}{\theta}{}^{o}_{m})^{2}\cdot\rho_{1}(\mathbf{m})<$$
$$<f^{-1}\cdot\mathbf{m}^{\mathrm{T}}(A)\,\mathbf{H}(A,S)\,\mathbf{m}(A)+f^{-1}\cdot q(\mathbf{m})\,,\qquad(84)$$

$\mathbf{H}(A,S)$ – $(n(A)\times n(A))$-матрица введена в (72); $\rho_1(\mathbf{m})$, f и $q(\mathbf{m})$ – скалярные величины определены в (69), (79) и (82) соответственно.

Исследование степени влияния скалярных величин $\rho_1(\mathbf{m})$ и $q(\mathbf{m})$ на условие редукции (84) в условиях схемы квазиповторных наблюдений, проведённое методом статистических испытаний, показало, что этими величинами можно пренебречь. Тогда условие редукции с учётом (79), принимает вид

$$(\overset{o}{\theta}{}^{o}_{m})^{2}\cdot\mathbf{m}^{\mathrm{T}}(A)\,\mathbf{\Sigma}_{\xi}^{-1}\,\mathbf{M}(A,S)\,\mathbf{m}(A)<1\,.\qquad(85)$$

Из (85) следует, что возможность редукции модели может быть обусловлена четырьмя причинами: а) малостью нормы коэффициента $\overset{o}{\theta}{}^{o}_{m}$; б) малостью нормы вектора наблюдений регрессора $\mathbf{m}(A)$; в) малым значением квадратичной формы $\mathbf{m}^{\mathrm{T}}(A)\,\mathbf{\Sigma}_{\xi}^{-1}\,\mathbf{M}(A,S)\,\mathbf{m}(A)$; г) малым объёмом выборок наблюдений $n(B)$.

Редукция модели, оптимальной по составу регрессоров, означает, что при выполнении соотношения между параметрами модели (85) следует исключить регрессор $\mathbf{m}$ из модели. Редуцированная (упрощённая) модель будет иметь меньшую ошибку прогнозирования выходной переменной на новых выборках по сравнению с моделью, построенной на истинной структуре.

Случай избыточного регрессора. Рассмотрим случай, когда в текущую структуру включен излишний регрессор. Предположим для простоты, что это регрессор с номером p, он является максимально удаленным предыдущим значением переменной и не участвует в формировании ее текущего значения, т. е. для структурных матриц выполняется

$$\mathbf{S}=[\,\overset{o}{\mathbf{S}}\,|\,\mathbf{s}\,],\qquad(86)$$

где $\overset{o}{\mathbf{S}}$ – структурная $(p\times\overset{o}{m})$-матрица истинной модели; $\mathbf{S}$ – структурная $(p\times(\overset{o}{m}+1))$-матрица текущей модели; $\mathbf{s}$ – $(p\times1)$-вектор

$$\mathbf{s}=\big(0,0,\ldots,0,1\big)^{\mathrm{T}}.\qquad(87)$$

Другими словами, в модели функционирования объекта (4) в формировании $\overset{*}{x}_i(h)$ не участвует величина $\overset{*}{x}_{i-p}(h)$, но в текущую модель она включена.

Используя (50)–(53), вычислим разность математических ожиданий критерия регулярности для текущей структуры S и истинной структуры $\overset{o}{S}$:

$$\Delta_2(S,\overset{o}{S})=E\{ARD^{*}(S)\}-E\{ARD^{*}(\overset{o}{S})\}=$$
$$=\Delta(B\,|\,A,S)+\mathrm{tr}\,[\,\mathbf{P}(B\,|\,A,S)\,]-\mathrm{tr}\,[\,\mathbf{P}(B\,|\,A,\overset{o}{S})\,]\,.\qquad(88)$$

В случае избыточного регрессора величина $\Delta(B\,|\,A,S)$ в (50) является нулевой. Для доказательства этого факта достаточно показать, что вектор смещения $\boldsymbol{\delta}(B\,|\,A,S)$ для случая избыточного регрессора является нулевым.

Действительно, различие в структурных матрицах (86)–(87) приводит к выполнению:

$$\mathbf{R}(S)=\mathbf{X}\mathbf{S}=\mathbf{X}[\overset{o}{\mathbf{S}}\,|\,\mathbf{s}]=[\mathbf{X}\overset{o}{\mathbf{S}}\,|\,\mathbf{X}\mathbf{s}]=[\mathbf{R}(\overset{o}{S})\,|\,\mathbf{r}\,], \tag{89}$$

где $\mathbf{r}$ – $(n\times 1)$-вектор наблюдений избыточного регрессора.

Для вектора смещения $\boldsymbol{\delta}(B\,|\,A,S)$ в случае избыточного регрессора выполняется

$$\boldsymbol{\delta}(B\,|\,A,S)=\overset{o}{\mathbf{y}}(B)-\mathbf{R}(B,S)\mathbf{C}(A,S)\overset{o}{\mathbf{y}}(A)=\mathbf{R}(B,\overset{o}{S})\overset{o}{\boldsymbol{\theta}}-\mathbf{R}(B,S)\mathbf{C}(A,S)\mathbf{R}(A,\overset{o}{S})\overset{o}{\boldsymbol{\theta}}=$$

$$=\mathbf{R}(B,\overset{o}{S})\overset{o}{\boldsymbol{\theta}}-[\,\mathbf{R}(B,\overset{o}{S})\ \vdots\ \mathbf{r}(B)]\times$$

$$\times\left(\begin{bmatrix}\mathbf{R}^{\mathrm{T}}(A,\overset{o}{S})\\ \hdashline \mathbf{r}^{\mathrm{T}}(A)\end{bmatrix}\boldsymbol{\Sigma}_{\xi}^{-1}[\,\mathbf{R}(A,\overset{o}{S})\ \vdots\ \mathbf{r}(A)]\right)^{-1}\begin{bmatrix}\mathbf{R}^{\mathrm{T}}(A,\overset{o}{S})\\ \hdashline \mathbf{r}^{\mathrm{T}}(A)\end{bmatrix}\boldsymbol{\Sigma}_{\xi}^{-1}\mathbf{R}(A,\overset{o}{S})\overset{o}{\boldsymbol{\theta}}. \tag{90}$$

Вычисляя аналогично (58)–(73), получаем $\boldsymbol{\delta}(B\,|\,A,S)=\mathbf{0}$ и $\Delta(B\,|\,A,S)=0$.

Для случая избыточного регрессора, аналогично (75)–(82), получаем

$$\mathrm{tr}\,[\,\mathbf{P}(B\,|\,A,S)-\mathbf{P}(B\,|\,A,\overset{o}{S})]=$$

$$=f^{-1}(\overset{o}{S})\cdot\mathbf{r}^{\mathrm{T}}(A)\mathbf{M}^{\mathrm{T}}(A,\overset{o}{S})\mathbf{M}(A,\overset{o}{S})\mathbf{r}(A)+f^{-1}(\overset{o}{S})\cdot q(\mathbf{r}), \tag{91}$$

$$f(\overset{o}{S})=\mathbf{r}^{\mathrm{T}}(A)\boldsymbol{\Sigma}_{\xi}^{-1}\mathbf{r}(A)-\mathbf{r}^{\mathrm{T}}(A)\boldsymbol{\Sigma}_{\xi}^{-1}\mathbf{R}(\overset{o}{S})(\mathbf{R}^{\mathrm{T}}(\overset{o}{S})\boldsymbol{\Sigma}_{\xi}^{-1}\mathbf{R}(\overset{o}{S}))^{-1}\mathbf{R}^{\mathrm{T}}(\overset{o}{S})\boldsymbol{\Sigma}_{\xi}^{-1}\mathbf{r}(A)=$$

$$=\mathbf{r}^{\mathrm{T}}(A)\boldsymbol{\Sigma}_{\xi}^{-1}\mathbf{M}(A,\overset{o}{S})\,\mathbf{r}(A) \tag{92}$$

– положительная величина, поскольку матрицы $\boldsymbol{\Sigma}_{\xi}^{-1}$ и $\mathbf{M}(A,\overset{o}{S})$ положительно определены; $\mathbf{M}(A,\overset{o}{S})$ – $(n\times n)$-матрица аналогична (67)–(68); $\mathbf{r}(A)$ – $(n\times 1)$-вектор, определённый в (89); для скалярной величины $q(\mathbf{r})$ выполняется

$$q(\mathbf{r})=g(\mathbf{r})-\mathbf{g}^{\mathrm{T}}(R)\mathbf{B}^{-1}\mathbf{R}^{\mathrm{T}}(A,\overset{o}{S})\boldsymbol{\Sigma}_{\xi}^{-1}\mathbf{r}(A)-\mathbf{r}^{\mathrm{T}}(A)\boldsymbol{\Sigma}_{\xi}^{-1}\mathbf{R}(A,\overset{o}{S})\mathbf{B}^{-1}\mathbf{g}(R)+$$

$$+\mathbf{r}^{\mathrm{T}}(A)\boldsymbol{\Sigma}_{\xi}^{-1}\mathbf{R}(A,\overset{o}{S})\mathbf{B}^{-1}\mathbf{G}(R)\mathbf{B}^{-1}\mathbf{R}^{\mathrm{T}}(A,\overset{o}{S})\boldsymbol{\Sigma}_{\xi}^{-1}\mathbf{r}(A). \tag{93}$$

Учитывая (91), пренебрегая влиянием $f^{-1}(\overset{o}{S})\cdot q(\mathbf{r})$, получаем

$$\Delta_2(S,\overset{o}{S})\approx f^{-1}(\overset{o}{S})\cdot\mathbf{r}^{\mathrm{T}}(A)\mathbf{H}(A,\overset{o}{S})\mathbf{r}(A)>0, \tag{94}$$

где матрица $\mathbf{H}(A,\overset{o}{S})=\mathbf{M}^{\mathrm{T}}(A,\overset{o}{S})\mathbf{M}(A,\overset{o}{S})$ – положительно определена, а величина $f^{-1}(\overset{o}{S})$ положительна, поэтому величина (94) положительна. Из (94) следует, что в случае избыточного регрессора истинная структура $\overset{o}{S}$ всегда лучше структуры S, а регрессор $\mathbf{r}$ действительно не следует включать в модель.

Заключение

По принципам метода группового учета аргументов построен и исследован критерий регулярности для моделирования в классе авторегрессионных уравнений в условиях квазиповторных наблюдений. Получены условия существования оптимального множества регрессоров, зависящие от коэффициентов моделей, матриц наблюдений регрессоров, дисперсии случайных составляющих и объемов выборок. Выявлены закономерности редукции (упрощения) оптимальной регрессионной модели. Редуцированная модель будет иметь меньшую ошибку прогнозирования выходной переменной на новых выборках по сравнению с моделью, построенной на истинной структуре.

Литература

1. Современные методы идентификации систем : пер. с англ. / Под ред. П. Эйкхоффа. – М. : Мир, 1983. – 400 с.
2. Льюнг Л. Идентификация систем. Теория для пользователя : пер. с англ. – М. : Наука, 1991. – 432 с.
3. Green W. H. Econometric Analysis. – Fifth edition. New Jersey : 2002. – 1056 p.
4. Кунцевич В.М. Управление в условиях неопределенности: гарантированные результаты в задачах управления и идентификации. – Киев : Наук. думка, 2006.– 264 с.
5. Markovsky I., Van Huffel S. Overview of total least squares methods // Signal Processing. – 2007. – 87. – P. 2283–2302.
6. Söderström T. Errors-in-variables methods in system identification // Automatica. – 2007. – 43 (6). – P. 939–958.
7. Губарев В. Ф., Гуммель А. В., Жуков А. О. Особенности и взаимосвязь задач идентификации и управления в условиях неопределенности // Международный научно-технический журнал “Проблемы управления и информатики”. – 2008. –№ 5. –С. 23–38.
8. Ивахненко А. Г. Индуктивный метод самоорганизации моделей сложных систем. – К. : Наук. думка, 1982. – 296 с.
9. Ивахненко А. Г., Степашко В. С. Помехоустойчивость моделирования. – К. : Наук. думка, 1985. – 216 с.
10. Ивахненко А. Г., Юрачковский Ю. П. Моделирование сложных систем по экспериментальным данным. – М. : Радио и связь, 1987. – 120 с.
11. Muller J.-A., Lemke F. Self-organizing Data Mining. Extracting Knowledge from Data. – Hamburg : Libri, 2000. – 250 p.
12. Сарычев А. П. Идентификация состояний структурно-неопределенных систем. – Днепропетровск : Институт технической механики, 2008. – 268 с.
13. Сарычев А. П. Идентификация параметров систем авторегрессионных уравнений при известных ковариационных матрицах // Международный научно-технический журнал “Проблемы управления и информатики”. – 2012. – № 3. – С. 14–30.
14. Сарычев А. П. Исследование методом статистических испытаний итерационной процедуры для идентификации параметров системы авторегрессионных уравнений // Системні технології. – Випуск 3 (92). – 2014. – С. 77–89.
15. Себер Дж. Линейный регрессионный анализ : пер. с англ. – М. : Мир, 1980. – 456 с.
16. Ермаков С. М., Жиглявский А. А. Математическая теория оптимального эксперимента. – М. : Наука, 1987. – 320 с.

Literatura

1. Sovremennye metody Identifikacii sistem: per. s angl. / Pod red. P. Eikhoff. – M. : Mir, 1983. – 400 s.
2. L’juing L. Identifikacija system. Teorija dlja pol’zovateleja : per. s angl. – M. : Nauka, 1991. – 432 s.
3. Green W. H. Econometric Analysis. – Fifth edition. New Jersey : 2002. – 1056 p.
4. Kuntsevich V. M. Upravlenie v uslovijakh neopredelennosti: garantirovannye rezul’taty v zadachakh upravlenija I identifikacii. – Kiev : Nauk. dumka, 2006.– 264 s.

5. Markovsky I., Van Huffel S. Overview of total least squares methods // Signal Processing. – 2007. – 87. – P. 2283–2302.
6. Söderström T. Errors-in-variables methods in system identification // Automatica. – 2007. – 43 (6). – P. 939–958.
7. Gubarev V. F., Gumel' A. V., Zhukov A. O. Osobennosti i vzaimosvjaz' zadach identifikacii i upravlenija v uslovijakh neopredelennosti // Problemy upravlenija i informatiki. – 2008. –№ 5. – S. 23–38.
8. Ivakhnenko A. G. Induktivnyj metod samoorganizacii modelej slozhnykh system / A.G Ivakhnenko. – K. : Naukova dumka, 1982. – 296 s.
9. Ivakhnenko A. G. Pomekhoustojchivost' modelirovanija / A.G Ivakhnenko, V. S. Stepashko. – K. : Naukova dumka, 1985. – 216 s.
10. Ivakhnenko A. G. Modelirovanie slozhnykh system po eksperimental'nym dannym / A.G Ivakhnenko, Uj. P. Ujrachkovskij. – M. : Radio I svjaz', 1987. – 120 s.
11. Muller J.-A., Lemke F. Self-organizing Data Mining. Extracting Knowledge from Data. – Hamburg : Libri, 2000. – 250 p.
12. Sarychev A. P. Identifikacija sostojanij strukturno-neopredelennykh system / A. P. Sarychev. – Dnepropetrovs'k : NAN Ukrainy I NKA Ukrainy, Institut tekhnichnoj mekhaniki, 2008. – 268 s.
13. Sarychev A. P. Identifikacija parametrov system avtoregressionnykh uravnenij pri izvestnykh rjvariacionnykh matricakh / A. P. Sarychev // Problemy upravlenija i informatiki. – 2012. – № 3. – S. 14–30.
14. Sarychev A. P. Issledovanie metodom statisticheskikh ispytanij iteracionnoj procedyry dlja Identifikacii parametrov systemy avtoregressionnykh uravnenij / A. P. Sarychev // Sistemni tekhnologij. – Vypusk 3 (92). – 2014. – S. 77–89.
15. Seber Dzh. Lineinyj pegressionnyj analiz : per. s angl. – M. : Mir, 1980. – 456 s.
16. Ermakov S. M., Zhigljavskij A. A. Matematicheskaja teoriya optimal'nogo eksperimenta. – M. : Nauka, 1987. – 320 s.

RESUME

A.P. Sarychev

Linear Autoregression Based on the Group Method of Data Handling in Conditions of Quasirepeated Observations

Determination of the order of autoregression model in conditions of structural uncertainty by quantity and structure of regressors is the important problem in the theory of identification. In the Group Method of Data Handling at modeling in the class of autoregression equations, the criterion of regularity is traditionally applied. Training sample is formed by the first $n(A)$ observations and testing sample is formed by the subsequent $n(B)$ observation, and it is carried out $n(A)+n(B)=n$, where n is volume of initial sample. Application of regularity criterion in such kind demands the control as dynamic properties of object can be shown unequally on different phases of transients. In this paper, the regularity criterion in the so-called scheme of quasirepeated observations that is possible in conditions of active experiment is offered. In this scheme the training (A) and testing (B) samples are received by special way as pair realizations of functioning of object with close initial conditions, qualitatively identical character of transients process and close conditions during the final moments of time.

It is proved, that the optimum set of regressors exists. The condition of a reduction of the optimum autoregression model is obtained. This condition depends on parameters of autoregression model and volumes of samples. The reduced model will have a smaller mistake of forecasting of a output variable on new samples of observations in comparison with model that constructed on true structure.

А.П. Сарычев

Линейная авторегрессия на основе метода группового учёта аргументов в условиях квазиповторных наблюдений

Определение порядка авторегрессионной модели в условиях структурной неопределённости по количеству и составу регрессоров является важной задачей в теории идентификации. В методе группового учёта аргументов при моделировании в классе авторегрессионных уравнений традиционно применяется критерий регулярности: обучающая выборка формируется первыми $n(A)$ наблюдениями временного ряда, проверочная выборка последующими $n(B)$ наблюдениями, причём выполняется $n(A)+n(B)=n$, где n – объём исходной выборки. Применение критерия регулярности в таком виде требует контроля, поскольку динамические свойства объекта могут проявляться неодинаково на разных фазах переходных процессов. В данной работе предлагается рассчитывать критерий регулярности в так называемой схеме квазиповторных наблюдений, которая возможна в условиях активного эксперимента. В этой схеме обучающая (A) и проверочная (B) выборки получены особым способом как пара реализаций функционирования объекта с близкими начальными условиями, качественно одинаковым характером переходных процессов и близкими состояниями в конечные моменты времени.

Доказано существование оптимального множества регрессоров. Получено условие редукции оптимальной авторегрессионной модели, которое зависит от параметров авторегрессионной модели и объёмов выборок. Редуцированная модель будет иметь меньшую ошибку прогнозирования выходной переменной на новых выборках наблюдений по сравнению с моделью, построенной на истинной структуре.

Поступила в редакцию 15.06.2015

